HOW TO MEASURE RECOVERIES AND PROVISIONS ON BANK LENDING

Methodology and Empirical Evidence

J. Dermine and C. Neto de Carvalho

Third draft: 28 March 2005

INSEAD (Fontainebleau) and Universidade Catolica Portuguesa (Lisbon), respectively. The authors are grateful to A. Resti, P. Jackson, M. Massa, J. Santos Silva, A. Sironi, M. Suominen, and J. Wahlen for helpful discussion, to J. Cropper for editorial assistance, and to
Introduction

The purpose of this chapter is to propose a methodology to estimate losses-given-default (LGD) and fair provisions on non-performing bank loans. The methodology is then applied to a unique set of micro-data on loans to small- and medium size enterprises (SMEs). LGD estimates are important inputs in the pricing of credit risk, and measurement of bank profitability and solvency. The international bank capital regulations, Basel II, require estimates of losses-given-default to calculate risk-weighted assets, and to estimate the fair level of provisions needed to adjust the amount of available capital.\(^1\)

The chapter is structured as follows. In Section 1, we argue that bank loans are likely to have some characteristics significantly different from those of corporate bonds. This justifies the need for specific studies on bank loans. In Section 2, the database on individual loans losses is presented. The mortality-based approach to analyze recovery rates on bad and doubtful loans is discussed in Section 3. Empirical evidence on cumulative recovery rates is presented in Section 4. In Section 5, we show how to use the methodology to calculate dynamic provisions over time. In Section 6, a multivariate statistical analysis of the determinants of loan losses-given-default and provisions is developed. Section 7 concludes the chapter.

Section 1. LGD Literature Review

If several authors have analyzed credit risk and losses-given-default on corporate bonds, few

\(^1\) Under the Basel II internal rating-based (IRB) approach, two cases must be distinguished: Estimates of losses-given-default are needed for the retail loan portfolio under the ‘foundation’ approach, while LGD on corporate loans are needed only for the ‘advanced’ approach. Banks using the IRB approach must compare their total eligible provisions with the total expected losses (Basel Committee, 2004).
studies have been applied to bank loans. The reason for this is that, as bank loans are private instruments, few data are publicly available.

The seminal work by Altman (1989) on US corporate bonds defaults was followed by studies on recovery rates, and on the degree of correlation between default frequencies and recovery rates (Frye 2000, a, b, 2003, Altman et al., 2003, or Acharya et al., 2003). These studies relied on publicly traded bond data. In view of this literature, one needs to justify a study on recovery on bank loans. Several arguments are proposed. The first two are that small firms are informationally more opaque, and that the relationship between the owner/manager of the firm and the bank is often very close (Allen et al., 2004). This has two implications. In the case of distress, the owner/manager has relatively more to lose because his or her skills will be firm-specific. Effort to repay a loan could be greater than in large public firms. Second, the close relationship with a bank might imply that the latter will hesitate to foreclose a loan, hoping to capture the option value of a future relationship (Dewenter and Hess, 2004). A third argument is that a bank might hesitate to foreclose large loans, if it believes that this can have local spillover effects on other firms, i.e., clients of the bank. In the arm’s-length transactional corporate bonds markets, these ‘option’ or ‘macro’ considerations will be ignored. A fourth reason is that a large number of distressed loans may create bottlenecks in the workout unit of the bank, with effect on recovery rates. Finally, in the case of distressed bank loans, knowledge of the timing of cash flow recoveries will be useful to measure interest rate risk, with computation of repricing gaps or duration measure, as well as to provide information to calculate dynamic loan loss provisions over time.

As mentioned above, few studies have focused on the bank loan markets because of the private nature of these transactions. Asarnow and Edwards (1995) examined 831 defaulted loans at Citibank over the period 1970-1993. They reported an average cumulative recovery rate of 65%. Carty and Lieberman (1996) measured the recovery rate on a sample of 58 bank loans. Based on secondary market prices for defaulted bank loans for the period 1989-1996, they reported an average defaulted bank loan price of 71%. The above studies focused on the US market. Hurt and Felsowayli (1998) analyzed 1,149 bank loan losses in Latin America over the period 1970-1996. They report an average recovery rate of 68.2 % and show that loan size is a contributory factor to loss rates, with large loan default exhibiting lower recovery rates. They attribute this to the fact that
large loans, often not secured, are made to economic groups that are family owned. None of the above studies provide information on the timing of recoveries.

Thanks to access to a unique data-base on loans to SMEs, this chapter provides empirical evidence on the timing of recovery on individual bad and doubtful loans, on cumulative recovery rates, and their economic determinants.

Section 2. Bank Loan Losses, Database and Measurement Issue

The database was provided by the largest private bank in Portugal, Banco Comercial Português (BCP). It consists of 374 non-performing loans granted to SMEs over the period June 1995 to December 2000.

Table 1 (Panels A and B) provides information on the number of defaults per year, and the amount of debt outstanding at the time of default. Table 1 (Panels C and D) shows the number of loans with guarantee or collateral, the age of the firm, and the number of years of relationship with the bank. One observes that the series of 374 default cases is distributed evenly over the six-year period, and that the distribution of debt outstanding at the time of default is highly skewed toward the low end. The available information in the database includes the history of the loan after a default has been identified, the type of collateral or guarantees, the industry classification, the interest rate charged on the loan, the internal rating attributed by the bank, the age of the firm, and the length of relationship with the bank.

In Table 1 (Panel C), the various forms of guarantees or collateral are reported. These include:

- Guarantees
- Real estate collateral
- Physical collateral (inventories)
- Financial collateral (bank deposits, bonds or shares)

---

2 An additional study includes La Porta, Lopez-de-Silanes and Zamarripa (2003), who analyze loan default and losses-given-default in Mexico in the context of ‘related lending’, that is, lending to shareholders or directors of the bank.

3 The data gathered for this study do not include any reference to the identity of the clients or any other information which, according to Portuguese banking law, cannot be disclosed. A more complete description of
Insert Table 1

In 35.6% of the cases, there is no guarantee or collateral. Guarantees, which are used in 58.3% of the cases, refer to written promises made by the guarantor (often the owner or the firm’s director) that allows the bank to collect the debt against the personal assets pledged by the guarantor. Collateral is used in 15% of the cases.

Panel D of Table 1 shows the age of the firm, and the number of years of relationship with the bank. Companies have, on average, a life of 17 years, with extremes going from 6 months to 121 years. The average relationship with the bank is six years.  

Any empirical study of credit risk raises two measurement issues. Which criterion should be used to define the time of a default event? Which method should be used to measure the recovery rate on a defaulted transaction?

The criterion used for the classification of a loan in the ‘default’ category is critical for a study on recovery rates, as a different classification would lead to different results. Three ‘default’ definitions are used in the literature:

i) A loan is classified as ‘doubtful’ as soon as “full payment appears to be questionable on the basis of the available information”.

ii) A loan is classified as ‘in distress’ as soon as a payment (interest and/or principal) has been missed.

iii) A loan is classified as ‘in default’ when a formal restructuring process or bankruptcy procedure is started.

In this study, because of data availability, we adopt the second definition, that is, a loan is classified as ‘in default’ as soon as a payment is missed.  

the database is available in Dermine and Neto de Carvalho (2005).

4 The relatively short average relationship is due to the fact that the bank was created in 1985, after the deregulation of the Portuguese banking system.

5 For the sake of comparison, the definition of default adopted by the Basel Committee is as follows (Basel Committee, 2004, p. 92): “A default is considered to have occurred with regard to a particular obligor when either or both of the two following events have taken place:

a) The bank considers that the obligor is unlikely to pay its credit obligations to the banking group in full, without recourse by the bank to actions such as realizing security (if held).
The second methodological issue relates to the measurement of recovery on defaulted loans. There are two methodologies:

i) The price of the loan at the default date, defined most frequently as the trading price one month after the default. This approach has been used in studies on recoveries on corporate bond defaults.

ii) The discounted value of future cash flows recovered after the default date.

As no market price data are readily available for defaulted bank loans in Portugal, the second methodology -the present value of actual recovered cash flows- is the only feasible alternative. This approach was adopted by Asarnow and Edwards (1995), Carty and Lieberman (1996), Hurt and Felsovalyi (1998), and La Porta et al. (2003). The present value of cash flows recovered on impaired loans allows us to measure the proportion of principal and interest that is recovered after the default date. This approach has the advantage that, if the loan is fully repaid, the present value of the actual cash flows recovered will be equal to the outstanding balance at the default date. It should be noted that this PV amount could differ from the price of the loan at a time of default, which would incorporate the expected cash flows and proper risk or liquidity premia.

Two approaches will be used to analyze the recovery rates, a univariate mortality-based approach, and a multivariate statistical analysis of the determinants of recovery.

Section 3. A Mortality-based Approach to Analyze Recovery Rates

Having access to the history of cash flows on these loans after default, we can study the time distribution of recovery. With reference to the studies by Altman (1989) and Altman and Suggitt (2000), we apply the mortality approach. It must be noted that the mortality approach was applied to measure the percentage of bonds or loans that defaulted \( n \) years after origination. The application of mortality to loan recovery rates is, to the best of our knowledge, novel. It examines the percentage of a bad and doubtful loan which is recovered \( n \) months after the default date. This

---

b) The obligor is past due more than 90 days on any material credit obligation to the banking group. Overdrafts will be considered as being past due once the customer has breached an advised limit or been advised of a limit smaller than current outstandings.”
methodology is appropriate because the population sample is changing over time. For some default loans, i.e., those of June 1995, we have a long recovery history (66 months), while for the recent ‘2000’ loans in default, we have an incomplete history of recovery.

To define the concepts used to measure loan recovery rate and provisions on impaired loans⁶, it is, for expository reasons, useful to refer to a simple example. Formal definitions of concepts are presented in the Appendix. Consider a loan of €100 that enters the ‘default’ category in December 2000. We track the subsequent payments on this loan, assuming, for expository convenience, that all payments take place at the end of the year, with a final payment in December 2003. The contractual interest rate is 10%.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Loan outstanding</td>
<td>100</td>
<td>110</td>
<td>66</td>
<td>44</td>
</tr>
<tr>
<td>(before cash payment)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cash payment</td>
<td>0</td>
<td>50</td>
<td>26</td>
<td>14</td>
</tr>
<tr>
<td>Loan balance</td>
<td>100</td>
<td>60</td>
<td>40</td>
<td>30</td>
</tr>
<tr>
<td>(after cash payment)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Let us define the *Marginal Recovery Rate* at December 2001, MRR₁, as the proportion of the outstanding loan in December 2001 that is being paid, one period (in the example, one year) after default:

\[
MRR_1 = \frac{\text{Cash flow paid}_1}{\text{Loan balance}_1} = \frac{50}{110} = \frac{50}{1.10} / 100 = \frac{5}{11}
\]

The marginal recovery rate can also be interpreted as the percentage repayment on the loan outstanding, in present value terms.

---

⁶ Provisions on performing loans would require a different methodology. A mark-to-market approach is developed in Dermine and Bissada (2002), and Dermine (2005).
Let us define the *Percentage Unpaid Loan Balance* after payment in December 2001, PULB\(_1\), as the proportion of the December 2001 loan balance that remains to be paid one period after default:

\[
PULB_1 = 1 - MRR_1 = 1 - \frac{5}{11} = \frac{6}{11}
\]

One can define the *Marginal Recovery Rate* and the *Percentage Unpaid Loan Balance* at December 2002, two periods after default:

\[
MRR_2 = \frac{\text{Cash flow paid}_2}{\text{Loan}_2} = \frac{26}{66}
\]

\[
PULB_2 = 1 - MRR_2 = 1 - \frac{26}{66} = \frac{40}{66}
\]

And, the *Cumulative Recovery Rate* in December 2002 for a loan defaulting in December 2000, CRR\(_{0,2}\), is defined as:

\[
CRR_{0,2} = 1 - (PULB_1 \times PULB_2)
\]

\[
= 1 - \left( \frac{6}{11} \times \frac{40}{66} \right) = 1 - \frac{240}{726} = \frac{81}{121}
\]

\[
= \left( \frac{81}{1.1^2} \right) / 100 = 66.9\%
\]

Similarly, one can calculate a *Marginal Recovery Rate*, MRR\(_3\), a *Percentage Unpaid Loan Balance*, PULB\(_3\), and a *Cumulative Recovery Rate*, CRR\(_{0,3}\), at December 2003:

\[
MRR_3 = \frac{\text{Cash flow paid}_3}{\text{Loan}_3}
\]

\[
= \frac{14}{44}
\]

\[
PULB_3 = 1 - MRR_3 = 1 - \frac{14}{44} = \frac{30}{44}
\]

\[
CRR_{0,3} = (1 - (PULB_1 \times PULB_2 \times PULB_3))
\]

\[
= 1 - \left( \frac{6}{11} \times \frac{40}{66} \times \frac{30}{44} \right) = 1 - \frac{7,200}{31,944} = 1 - \frac{30}{133.1}
\]

\[
= \frac{103.1}{133.1} = \left( \frac{103.1}{1.1^3} \right) / 100 = 77.5\%
\]

Similar to Carey (1998), Asarnow and Edwards (1995), Carty and Lieberman (1996), Hurt and Felsovalyi (1998), and La Porta *et al.* (2003), the *Cumulative Recovery Rate* at time \( T \) on a loan
balance outstanding at time 0, CRR\(_{0,T}\), represents the proportion of the initial defaulted loan that has been repaid\(^7\) (in present value terms), \(T\) periods after default.

Finally, the *Loan-Loss Provision*, LLP\(_0\), on a loan balance outstanding at the default date, December 2000, is defined as:

\[
LLP_0 = 1 - CRR_{0,3} = 1 - 0.775 = 22.5\%.
\]

This percentage, 22.5\%, represents the percentage of the loan that will not be recovered (interest included) in the future. With perfect foresight, it would serve as a base for provisioning on the loan at the time of default.

A related question is to analyze the provision on a loan, \(n\) years after the default date. This allows the calculation of *dynamic* provisions. Moving forward, one can define a *Cumulative Recovery Rate*, CRR\(_{1,3}\), on a loan balance outstanding at December 2001, and a dynamic *Loan-Loss Provision*, LLP\(_1\), on a loan balance outstanding at December 2001. These are defined as:

\[
CRR_{1,3} = (1 - \frac{PULB_2}{66} \times \frac{PULB_3}{44}) = 1 - \frac{40}{66} \times \frac{30}{44} = 58.7\%
\]

\[
LLP_1 = 1 - CRR_{1,3} = 41.3\%
\]

In a similar manner, the *Cumulative Recovery Rate*, CRR\(_{2,3}\), and the dynamic *Loan-Loss Provision*, LLP\(_2\), on a loan balance outstanding at December 2002 are defined as:

\[
CRR_{2,3} = (1 - \frac{PULB_3}{64}) = 1 - \frac{30}{44} = 31.8\%
\]

\[
LLP_2 = 1 - CRR_{2,3} = 68.2\%
\]

Knowledge of marginal recovery rates over time thus allows one to calculate the dynamic evolution of provisions over time.

Having computed the cumulative recovery rate on individual loans, one can compute an arithmetic average cumulative recovery rate for the sample of loans. Alternatively, one can compute a principal-weighted average recovery rate that will take into account the size of each loan. A

---

\(^7\) One observe that 103.1, in CRR\(_{0,3}\), is the future capitalized value of interim cash flow recovered (103.1 = (50x1.1\(^2\)) + (26x1.1) + 14).
comparison of the sample weighted cumulative recovery rate with the average of recovery rates on individual loans will be indicative of a size effect.

Section 4. Cumulative Recovery Rates, Empirical Results

The sample marginal and cumulative recovery rates for the sample $T$ periods after a default date, $SMRR_T$ and $SCRR_T$, respectively, are reproduced in Figures 1 and 2.

In Figure 1, one observes that most of the marginal recovery rates in excess of 5% occur in the first five months after the default, and that the cumulative average recovery rate is almost completed after 48 months. As indicated in Figure 2, the unweighted (weighted) cumulative recovery rates after 36 and 48 months are, respectively, 67.3% (53.5%) and 70% (56.3%). The differences between the unweighted and weighted average cumulative recovery rates are an indicator that recovery on large loans is significantly lower.

It is also of interest to analyze the distribution of cumulative recovery rates across the sample of loans. The distribution of cumulative recovery rates after 48 months is reproduced in Figure 3. This figure shows a bi-modal distribution with many observations with low recovery, and many with complete recovery. These results are quite similar to those reported by Asarnow and Edwards (1995) and Schuermann (2004) for the US, and Hurt and Felsovalyi (1998) for Latin America. Loan portfolio models which incorporate a probability distribution for recovery rates should take into account this bi-modal distribution.

Section 5. Loan-Loss Provisioning, Empirical Evidence with a Mortality-

---

8 Note that, as is the case of other studies, these are gross recovery rates. Estimates of the bank’s internal recovery costs, which include the cost of the workout units and legal and accounting costs, are discussed in Dermine and Neto de Carvalho (2005).

9 Current commercial credit portfolio models do not incorporate the bi-modal distribution. For instance, CreditRisk+ assumes a fixed expected recovery rate within each band, while CreditMetrics uses a beta distribution (Crouhy et al., 2000).
based Approach

The accounting and finance literatures have analyzed three main provisioning issues related to private information held by banks on loans: the extent of earnings and capital smoothing (Hasan and Wall, 2004), the impact of reported provisions on a bank’s stock returns (Wahlen, 1994), and the systemic impact on the banking industry of disclosure on loan provisions by one bank (Grammatikos and Saunders, 1990). These studies are concerned with fair level of provision at the portfolio level. To the best of our knowledge, none have dealt with provisions at the micro level. As defined in Section 3, loan-loss provision is the complement of the cumulative recovery rate.

Insert Table 2

Cumulative recovery on outstanding balances on bad and doubtful loans are reported in Table 2. These dynamic cumulative recoveries are calculated on loan balances outstanding \( n \) months after the default date, with \( n \) varying from 0, the time of financial distress, to 37 months after the default date. Over time, the weighted (and unweighted) cumulative recovery rates on the unpaid loan balance are decreasing monotonously over time, 63.8% (72.6%) in month 0, to 23.3% (17.1%) in month 37. More complete information is obtained (Table 2), by dividing the sample into three groups, according to the absence of guarantee/collateral, the existence of guarantee, or the existence of collateral support. As expected, the cumulative recovery rates (unweighted, so as not to be affected by the size effect) is the highest for the case of loans with collateral for every period, for instance 85.3% at time 0, the default date. A counterintuitive and significant observation is that the (unweighted) cumulative recovery rates on loans without any type of guarantee or collateral dominates, in every period, the cumulative recovery rate on loans with guarantees. For instance, at time 0, (unweighted) recoveries on loans without any type of guarantee/collateral is 80.6%, vs.63.9% for recoveries on loans with guarantee (or 47.0% vs. 20.8%, 19 months after the default date). This result is likely to be due to two factors. First, as explained by the bank, guarantee or collateral support is not usually requested from reliable clients, so that the existence of a guarantee is an indicator of greater risk. Second, some borrowers are able to shift ownership of personal assets to other persons, so that, when the bank tries to execute the debt, there is not much left. This empirical result, if confirmed in other studies, implies that regulatory provisions should not necessarily penalize loans without guarantee, as the absence of a guarantee might be justified by higher expected recoveries.
Section 6. Loan-Loss Provisioning, a Multivariate Analysis

In the previous sections, univariate mortality-based estimates of cumulative recoveries and provisions were provided. In this section, we attempt to estimate empirically the determinants of the cumulative recovery rates and dynamic provisions. A discussion of the choice of explanatory variables and the econometric specification is followed by the empirical results.

Explanatory variables and econometric specification

Explanatory variables include the size of the loan, the type of guarantee/collateral support, past cumulative recovery, the age of the firm, and the industry sector. Fifteen business sectors have been created, with reference to the European Union’s NACE economic activity codes. Further aggregation, used in the econometric tests, leads to four activity sectors: real sector (activities with well identified real assets, such as land, mines or real estate property, which could be used for security), manufacturing, trade, and services. The size of the loan is included because some empirical studies and the sample univariate weighted and unweighted average cumulative recovery data have pointed out the effect of the loan size. Past cumulative collection is included on the assumption that a good level of collection could indicate a genuine effort by the borrower to repay the loan fully. A year dummy is included so as to have a better understanding of the volatility of the recovery rate over time. Finally, it is of interest to know the impact of guarantee/collateral, as, if statistically significant, this variable will be taken into account in calculating loan-loss provisions on bad and doubtful loans.10

The size of loan, past recovery rate, age of the firm, and length of relationship excepted, the explanatory variables will be represented by ‘dummy’ variables. The dependent variable, the cumulative loan recovery rate, is a continuous variable over the interval [0-1]. Due to the boundaries of the dependent variable, one cannot use the ordinary least square (OLS) regression,

10 Additional explanatory variables have also been tested, but found not significant over this sample: the number of years of the client’s relationship with the bank, the annual GDP rate of growth, the frequency of default in
\[ E(y^* x) = \beta_1 + \beta_2 x_2 + \ldots + \beta_k x_k = x \beta \] \hspace{1cm} (1)

as it cannot guarantee that the predicted values from the model will lie in the bounded interval (Greene, 1993). A common econometric technique is to use a transformation \( G(y) \) that maps the \([0-1]\) interval onto the whole real line \([-?, +?]\). There are several possible functional forms, but the most common ones are the cumulative normal distribution, the logistic function, and the log-log function. The log-log function is defined as:

\[ G(x \beta) = e^{-e^{-\beta}} \] \hspace{1cm} (2)

The cumulative normal distribution and the logistic function are symmetrically distributed, while the log-log function is asymmetric. This might be more appropriate with our data, since there is a significant concentration of observations near the extreme value ‘1’. Following up on Papke and Wooldridge (1996), the non-linear estimation procedure maximizes a Bernoulli log-likelihood function:

\[ l_i(\beta) = y_i [\log G(x_i \beta)] + (1- y_i) \log [1 - G(x_i \beta)] \] \hspace{1cm} (3)

**Empirical Results**

The log-log function has been estimated for the cumulative recovery rates (recovery measured until month 48)\(^{11}\) on balances outstanding at time 0, 4, 7, 13, 19, 25, and 37. The results are reported in Table 3.

Insert Table 3

The results confirm those of the univariate mortality-based approach, and add new insights. First the coefficient of the guarantee is negative and statistically significant at most time horizons. That is, losses and provisions on loans with guarantee are significantly higher than on loans with no

---

\(^{11}\) Because of data limitation (five years), the cumulative recovery is calculated up to 48 months after default. This does not seem too restrictive as Figure 2 indicates that most of the recovery is achieved 48 months after default. The data include the loans of 1995 and 1996, for which a 48-month recovery history is available.
guarantee/collateral. They should not receive preferential treatment in terms of provisioning. Second, the coefficient of collateral is positive at all time intervals, although statistically significant at two horizons only, 0 and 4 months. Third, and most interestingly, is the sign of the past recovery variable. It is positive and statistically significant at all horizons, 4 to 37 months, indicating that past recovery is a good indicator of future recovery. Finally, the loan size has a negative impact on recovery at time 0, confirming the observed univariate differences between unweighted and weighted recoveries. The loan size variable was not included in the other regressions on later recoveries, as it was found to be negatively correlated with the past recovery variable.

Section 7. Conclusion

A methodology developed to estimate bank loan losses-given-default and dynamic provisions, has been applied to a sample of 374 corporate distressed loans over the period 1995 to 2000. The estimates were based on the discounted value of cash flows recovered after the default event. A univariate mortality-based approach was first applied. The average cumulative recovery estimate of 71% is in the same order as that obtained in a few US studies. Partitioning of the sample indicates the impact on provisioning of collateral and guarantees. A multivariate approach was then applied to analyse the determinants of recovery rates. Three main conclusions can be drawn from this empirical study. The first is that the frequency distribution of loan losses-given-default appears bi-modal, with many cases presenting 0% recovery and other cases presenting a 100% recovery. Loan portfolio models should capture this characteristic. The second conclusion is that a multivariate analysis of the determinants of loan losses allows us to identify several statistically significant explanatory variables. These include the size of the loan, collateral, industry sector, year dummies, the age of the firm, and past recovery. Third, an analysis of cash flow recovery over time allows us to calculate a dynamic provisioning schedule. A word of caution on the empirical results is that this study, being based on a
dataset of a single bank, could capture some of the bank’s idiosyncracies.
Appendix: Recovery Concepts

For an individual loan $i$ in default, we define four concepts, $t$ denoting the number of periods after the initial default date $0$:

$$MRR_{i,t} = \text{Marginal Recovery Rate in period } t$$
$$= \frac{\text{Cash flow}_i \text{ paid at the end of period } t}{\text{Loan}_i \text{ outstanding at time } t}$$

$$PULB_{i,t} = \text{Percentage Unpaid Loan Balance at the end of period } t = 1 - MRR_{i,t}$$

The Cumulative Recovery Rate evaluated from the default date $0$ until infinity, $CRR_{i,0,\infty}$, and the Loan Loss Provision, $LLP_{i,0}$, are equal to:

$$CRR_{i,0,\infty} = \text{Cumulative Recovery Rate }_{\infty} \text{ periods after the default} = 1 - \prod_{t=1}^{\infty} PULB_{i,t}$$

$$LLP_{i,0} = \text{Loan-loss provisions} = 1 - CRR_{i,0,\infty}$$

For the sake of presentation, the loan-loss provision was calculated on a loan balance outstanding at the default date, $0$. In a more general dynamic provisions setting, the provision can be calculated on a loan balance outstanding at any date $n$ after the default date. For instance, one can compute the cumulative recovery on loan balances outstanding 4 or 13 months after the default date.
Table 1: Descriptive Statistics for the Sample of Bad and Doubtful Loans

<table>
<thead>
<tr>
<th>Panel A: Number of Defaults per year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
</tr>
<tr>
<td>1996</td>
</tr>
<tr>
<td>1997</td>
</tr>
<tr>
<td>1998</td>
</tr>
<tr>
<td>1999</td>
</tr>
<tr>
<td>2000</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Debt Outstanding at the Time of Default (€)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Observations</td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>0&lt;Debt&lt;50,000</td>
</tr>
<tr>
<td>50,000&lt;Debt&lt;100,000</td>
</tr>
<tr>
<td>100,000&lt;Debt&lt;150,000</td>
</tr>
<tr>
<td>150,000&lt;Debt&lt;200,000</td>
</tr>
<tr>
<td>200,000&lt;Debt&lt;250,000</td>
</tr>
<tr>
<td>250,000&lt;Debt&lt;300,000</td>
</tr>
<tr>
<td>300,000&lt;Debt&lt;350,000</td>
</tr>
<tr>
<td>Debt&gt; 350,000</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: Number of Loans with Guarantee or Collateral</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Observations</td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>No Guarantee/Collateral</td>
</tr>
<tr>
<td>Guarantee</td>
</tr>
<tr>
<td>Real Estate Collateral</td>
</tr>
<tr>
<td>Physical Collateral</td>
</tr>
<tr>
<td>Financial Collateral</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel D: Age of Borrowing Firm and Age of Relationship with the Bank (Years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Age of Borrowing Firm</td>
</tr>
<tr>
<td>Age of Relationship</td>
</tr>
</tbody>
</table>
Table 2: Mortality-based estimate of cumulative recoveries on loan balances outstanding \( n \) months after default

<table>
<thead>
<tr>
<th>Number of months after default (( n ))</th>
<th>Total sample</th>
<th>No Collateral / No Guarantee</th>
<th>Guarantee only</th>
<th>Collateral with or without guarantee</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unweighted</td>
<td>Weighted</td>
<td>Unweighted</td>
<td>Weighted</td>
</tr>
<tr>
<td>0</td>
<td>72.6%</td>
<td>63.8 %</td>
<td>80.6%</td>
<td>80.0%</td>
</tr>
<tr>
<td>4</td>
<td>57.8%</td>
<td>55.7%</td>
<td>69.6%</td>
<td>75.3%</td>
</tr>
<tr>
<td>7</td>
<td>50.2%</td>
<td>47.7%</td>
<td>63.3%</td>
<td>70.8%</td>
</tr>
<tr>
<td>13</td>
<td>41.0%</td>
<td>42.1%</td>
<td>56.6%</td>
<td>66.7%</td>
</tr>
<tr>
<td>19</td>
<td>33.0%</td>
<td>35.8%</td>
<td>47.0%</td>
<td>59.4%</td>
</tr>
<tr>
<td>25</td>
<td>27.1%</td>
<td>32.2%</td>
<td>41.3%</td>
<td>56.7%</td>
</tr>
<tr>
<td>37</td>
<td>17.1%</td>
<td>23.3%</td>
<td>28.3%</td>
<td>43.2%</td>
</tr>
</tbody>
</table>

Note: These are cumulative recovery rates up to month 68, calculated on the loan balance outstanding \( n \) months after the default date. The unweighted average recovery is the arithmetic average of recoveries on individual loans. Weighted average recoveries are weighted by loan balances.
Table 3: Statistical estimate of cumulative recoveries on loan balances outstanding \( n \) months after default

(The estimate of the parameter is followed by the p-value)

<table>
<thead>
<tr>
<th>Number of Months after Default ( (n) )</th>
<th>0</th>
<th>4</th>
<th>7</th>
<th>13</th>
<th>19</th>
<th>25</th>
<th>37</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.79 (0.00* )</td>
<td>1.08 (0.02* )</td>
<td>1.33 (0.01* )</td>
<td>0.93 (0.07 )</td>
<td>0.63 (0.21 )</td>
<td>0.66 (0.2 )</td>
<td>-0.28 (0.53 )</td>
</tr>
<tr>
<td>Past Recovery</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Loan Size</td>
<td>-0.77(0.00* )</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Year 1996</td>
<td>0.31 (0.24 )</td>
<td>0.11(0.65 )</td>
<td>-0.28 (0.24 )</td>
<td>-0.16 (0.50 )</td>
<td>-0.07 (0.75 )</td>
<td>-0.15 (0.48 )</td>
<td>-0.5 (0.04* )</td>
</tr>
<tr>
<td>Guarantee</td>
<td>-0.38(0.15 )</td>
<td>-0.43 (0.08 )</td>
<td>-0.65 (0.01 )</td>
<td>-0.75 (0.00 )</td>
<td>-0.63 (0.01* )</td>
<td>-0.63 (0.01* )</td>
<td>-0.80 (0.002* )</td>
</tr>
<tr>
<td>Collateral</td>
<td>1.72(0.00* )</td>
<td>0.87 (0.01* )</td>
<td>0.63 (0.17 )</td>
<td>0.45 (0.30 )</td>
<td>0.39 (0.39 )</td>
<td>0.20 (0.68 )</td>
<td>0.85 (0.18 )</td>
</tr>
<tr>
<td>II.Manufacturing</td>
<td>-1.2(0.02* )</td>
<td>-1.08 (0.02* )</td>
<td>-0.91 (0.05* )</td>
<td>-0.84 (0.08 )</td>
<td>-0.90(0.07 )</td>
<td>-1.06 (0.04 )</td>
<td>-0.69 (0.2 )</td>
</tr>
<tr>
<td>III.Trade sector</td>
<td>-1.25(0.01* )</td>
<td>-1.17 (0.01* )</td>
<td>-1.18 (0.01* )</td>
<td>-0.99 (0.04* )</td>
<td>-0.96 (0.05 )</td>
<td>-0.95(0.3 )</td>
<td>-0.30 (0.55 )</td>
</tr>
<tr>
<td>IV.Services</td>
<td>-0.42(0.48 )</td>
<td>-0.59 (0.3 )</td>
<td>-0.37 (0.53 )</td>
<td>-0.12 (0.83 )</td>
<td>-0.17 (0.77 )</td>
<td>-0.77 (0.17 )</td>
<td>0.14 (0.8 )</td>
</tr>
<tr>
<td>Age of firm</td>
<td>0.01 (0.04* )</td>
<td>0.001(0.06 )</td>
<td>-0.0007(0.39 )</td>
<td>-0.0002(0.88 )</td>
<td>0.00(0.76 )</td>
<td>-0.0007(0.3 )</td>
<td>-0.0006(0.20 )</td>
</tr>
<tr>
<td>Pseudo R(^2)</td>
<td>0.18</td>
<td>0.36</td>
<td>0.35</td>
<td>0.41</td>
<td>0.43</td>
<td>0.44</td>
<td>0.52</td>
</tr>
<tr>
<td>Wald (Qui-squared) test (p-value)</td>
<td>43.4</td>
<td>65.1</td>
<td>43.9</td>
<td>55.9</td>
<td>54.8</td>
<td>36</td>
<td>33</td>
</tr>
<tr>
<td>Reset</td>
<td>0.23 (0.82* )</td>
<td>-0.39 (0.69* )</td>
<td>-0.79 (0.43* )</td>
<td>0.22 (0.83* )</td>
<td>0.95 (0.34* )</td>
<td>1.0 (0.31* )</td>
<td>0.13(0.89* )</td>
</tr>
<tr>
<td>Number of observations</td>
<td>153</td>
<td>125</td>
<td>112</td>
<td>96</td>
<td>85</td>
<td>77</td>
<td>66</td>
</tr>
</tbody>
</table>

* Significant at the 5% level.
Note: The table presents the result of the log-log regression. The dependent variable is the cumulative recovery rate on loan balances outstanding at 4, 7, 13, 19, 25, and 37 months after the default date. The explanatory variables include a constant term, the percentage for the initial loan already recovered, the loan size, a dummy for guarantee, and collateral, and an industry sector. With reference to European Union economic activities codes (NACE), 15 sectors are first defined as follows: Sector 1 (Agriculture): 1, 2, 5, 20; Sector 2 (Mining): 11, 13, 14; Sector 3 (Construction): 45; Sector 4 (Hotel, restaurant): 55; Sector 5 (Real estate): 70; Sector 6 (Food-beverages): 15, 16; Sector 7 (Textiles): 17, 18, 19; Sector 8 (Chemicals): 23, 24, 25; Sector 9 (Machinery): 26 to 37; Sector 10 (Paper, printing): 21, 22; Sector 11 (Other mineral; cement): 26; Sector 12 (Wholesale trade): 50, 51; Sector 13 (Retail trade): 52; Sector 14 (Transport): 60 to 64; Sector 15. (Other services): 71 to 93. These are then aggregated into four sectors: Real (sectors 1 to 5); Manufacturing (sectors 5 to 11); Trade, (sectors 12+13); Services (sectors 14+15).

Because of data availability, the cumulative recovery is calculated up to 48 months after the default date. This is unlikely to create a bias, as Figure 2 indicates that most of the recovery is achieved 48 months after the default date.
Figure 1: Unweighted Marginal Recovery Rate at time $t + n$ (MRR$_{t+n}$)

Note: This figure presents the marginal recovery $n$-months after default. The mortality-based approach is used to calculate the marginal recoveries.
Figure 2: Unweighted and Weighted Cumulative Recovery Rate at time $t + n$ (CRR$_{t+n}$)

![Diagram showing cumulative recovery rates over time](image)

Notes: The figure presents the cumulative weighted and unweighted recovery rates $n$- months after default. They have been calculated with the mortality-based approach.

Figure 3: Distribution of Cumulative Recovery Rates forty-eight Months after Default. The sample includes the loans of 1995 and 1996, for which 48-month recovery data are available.

![Diagram showing loan recovery rates distribution](image)
References:


Abstract

Estimates of loan losses-given-default are essential parameters to price credit risk, to calculate a fair level of provision, and to evaluate bank profitability and solvency. The empirical literature on credit risk has relied mostly on the corporate bond market to estimate losses in the event of default. The reason for this is that, as bank loans are private instruments, few micro data on loan losses and cash flow recovery are publicly available. The main contribution of this chapter is to provide a mortality-based methodology to calculate loan losses-given-default and provisioning on impaired bank loans. The methodology is then applied to a unique set of micro-data on distressed loans of a European bank.